U.S. Department of Health and Human Services

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

A National Web Conference on Improving Health IT Safety Through the Use of Natural Language Processing to Improve Accuracy of EHR Documentation

Presented by:
Thomas Payne, M.D.
Li Zhou, M.D., Ph.D.

Moderated by:
Chris Dymek, Ed.D.
Agency for Healthcare Research and Quality

February 7, 2017

Agenda

- Welcome and Introductions
- Presentations
- Q&A Session With Presenters
- Instructions for Obtaining CME Credits

Note: After today's Webinar, a copy of the slides will be emailed to all participants.

AHRQ's Mission

To produce evidence to make health care safer, higher quality, more accessible, equitable, and affordable, and work within the U.S. Department of Health and Human Services and with other partners to make sure that the evidence is understood and used.

How AHRQ Makes a Difference

- AHRQ **invests in research and evidence** to understand how to make health care safer and improve quality.
- AHRQ creates materials to **teach and train** health care systems and professionals to **catalyze** improvements in care.
- AHRQ **generates measures and data** used to track and improve performance and evaluate progress of the U.S. health system.

AHRQ Health IT Funding

Apply now for **Research Demonstration and Dissemination Projects** in clinical decision support:

- Scale and spread existing clinical decision support for patient-centered outcomes research
  http://grants.nih.gov/grants/guide/pa-files/PA-16-283.html

- Develop new clinical decision support for patient-centered outcomes research
  http://grants.nih.gov/grants/guide/pa-files/PA-16-282.html

The Division of Health IT is **actively seeking R01, R03, R18, and R21 applications** to study:

- Design, implementation, usability, and safe use of health IT
  http://grants.nih.gov/grants/guide/notice-files/NOT-HS-16-009.html

- Use of health IT for patient-reported outcomes to improve quality
  http://grants.nih.gov/grants/guide/notice-files/NOT-HS-16-015.html

- Utilizing Health Information Technology to Scale and Spread Successful Practice Models Using Patient-reported Outcomes
  https://grants.nih.gov/grants/guide/pa-files/PA-17-077.html

Presenter and Moderator Disclosures

The following presenters and moderator have no financial interests to disclose:

- Thomas Payne, M.D.
- Li Zhou, M.D., Ph.D.
- Chris Dymek, Ed.D.

This continuing education activity is managed and accredited by the Professional Education Services Group (PESG), in cooperation with AHRQ, AFYA, and RTI.

PESG, AHRQ, AFYA, and RTI staff have no financial interests to disclose.

Commercial support was not received for this activity.

Presenter Grant Support

Recent grant support for our presenters includes:

Dr. Thomas Payne:

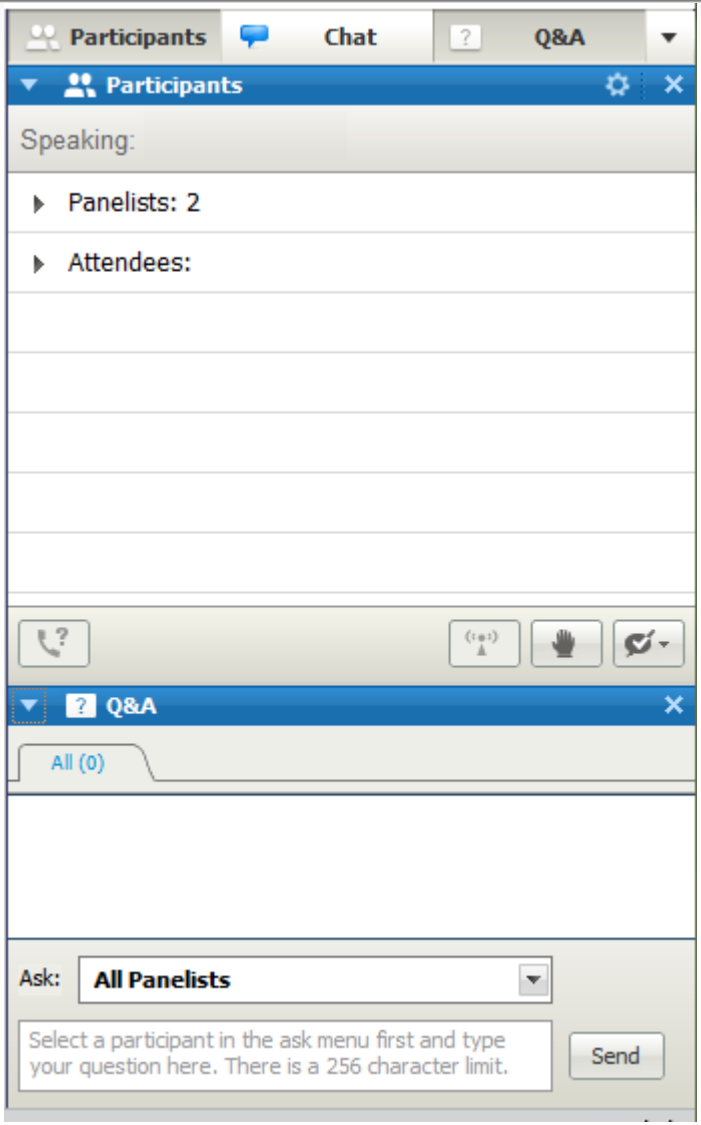
AHRQ HS023631
AUR A100077

Dr. Li Zhou

AHRQ HS024264
AHRQ HS022728
Controlled Risk Insurance Company (CRICO)
Brigham Care Redesign Incubator and Startup Program

How to Submit a Question

- At any time during the presentation, type your question into the “Q&A” section of your WebEx Q&A panel.
- Please address your questions to “All Panelists” in the drop-down menu.
- Select “Send” to submit your question to the moderator.
- Questions will be read aloud by the moderator.



Learning Objectives

At the conclusion of this activity, the participant will be able to do the following:

1) Discuss the development and evaluation of an enhanced electronic note system that leverages voice recognition and Natural Language Processing (NLP) technologies to create electronic physician notes in the EHR.

2) Discuss the challenges of introducing speech recognition technology into existing medical culture and current clinician workflow, including user preferences and the quality of documents generated by this technology.

3) Explain the need for an automated error detection system using NLP for improving the accuracy and quality of speech recognition generated medical documents, and discuss the development and evaluation of such a system.

Improving Accuracy of Electronic Notes Using a Faster, Simpler Approach

Thomas H. Payne, M.D.
University of Washington

Objectives

- Review problems with current physician note-writing practices.
- See how these problems might be addressed using current technologies and commercial EHRs.
- Understand barriers to changing physician documentation practices and how to address them.

Disclaimer

- Analysis still underway
- Statistical testing not yet performed

Final results may differ from those presented here.

Terms Used in This Webinar

Automatic speech recognition: Using software to convert spoken speech into text.

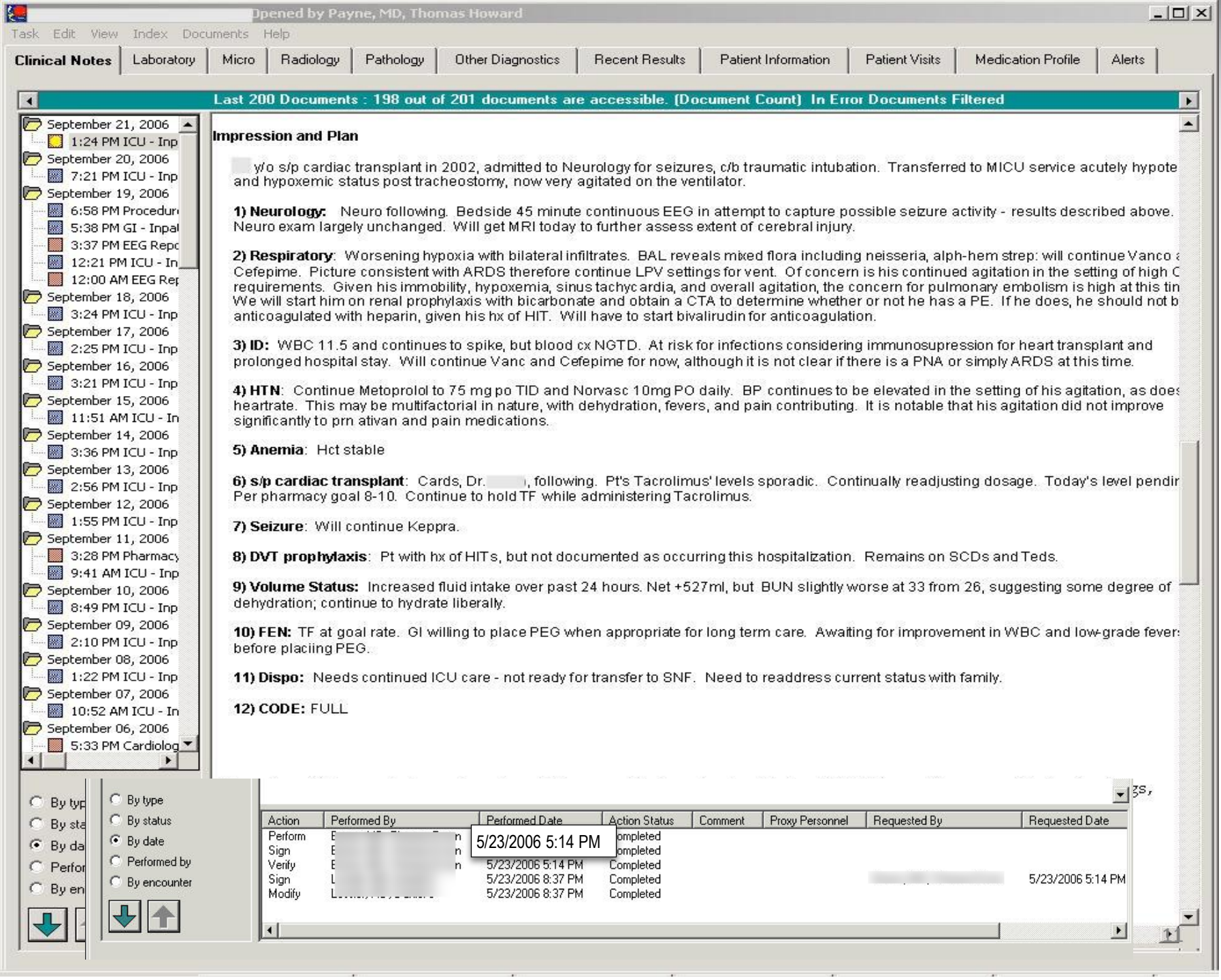
Ways to use automatic speech recognition software:

- ***Interactive*:** The user speaks into a microphone and watches the screen as the voice is converted to text and the user corrects errors interactively.
- ***Noninteractive*:** The user creates a voice file containing the entire document by speaking into a telephone or voice recorder. Automatic speech recognition software converts the voice file to text in the background while the user is engaged in other activities. A transcriptionist or the user corrects errors.

What Hath We Wrought?

Other Problems With Notes

- Note bloat.
- Copy and paste is common, usually as a result of efforts to save time.
- Progress notes are finished so late that other team members may not see them until the next day.


Opened by Payne, MD, Thomas Howard
Task Edit View Index Documents Help
Clinical Notes
Laboratory
Micro
Radiology
Pathology
Other Diagnostics
Recent Results
Patient Information
Patient Visits
Medication Profile
Alerts
Last 200 Documents : 198 out of 201 documents are accessible. (Document Count) In Error Documents Filtered
September 21, 2006
1:24 PM ICU - Inp
September 20, 2006
7:21 PM ICU - Inp
September 19, 2006
6:58 PM Procedur
5:38 PM GI - Inpal
3:37 PM EEG Repo
12:21 PM ICU - In
12:00 AM EEG Rep
September 18, 2006
3:24 PM ICU - Inp
September 17, 2006
2:25 PM ICU - Inp
September 16, 2006
3:21 PM ICU - Inp
September 15, 2006
11:51 AM ICU - In
September 14, 2006
3:36 PM ICU - Inp
September 13, 2006
2:56 PM ICU - Inp
September 12, 2006
1:55 PM ICU - Inp
September 11, 2006
3:28 PM Pharmacy
9:41 AM ICU - Inp
September 10, 2006
8:49 PM ICU - Inp
September 09, 2006
2:10 PM ICU - Inp
September 08, 2006
1:22 PM ICU - Inp
September 07, 2006
10:52 AM ICU - In
September 06, 2006
5:33 PM Cardiolog
Impression and Plan
y/o s/p cardiac transplant in 2002, admitted to Neurology for seizures, c/b traumatic intubation. Transferred to MICU service acutely hypote and hypoxemic status post tracheostomy, now very agitated on the ventilator.
1) Neurology: Neuro following. Bedside 45 minute continuous EEG in attempt to capture possible seizure activity - results described above. Neuro exam largely unchanged. Will get MRI today to further assess extent of cerebral injury.
2) Respiratory: Worsening hypoxia with bilateral infiltrates. BAL reveals mixed flora including neisseria, alph-hem strep: will continue Vanco a Cefepime. Picture consistent with ARDS therefore continue LPV settings for vent. Of concern is his continued agitation in the setting of high C requirements. Given his immobility, hypoxemia, sinus tachycardia, and overall agitation, the concern for pulmonary embolism is high at this tin We will start him on renal prophylaxis with bicarbonate and obtain a CTA to determine whether or not he has a PE. If he does, he should not b anticoagulated with heparin, given his hx of HIT. Will have to start bivalirudin for anticoagulation.
3) ID: WBC 11.5 and continues to spike, but blood cx NGTD. At risk for infections considering immunosupression for heart transplant and prolonged hospital stay. Will continue Vanc and Cefepime for now, although it is not clear if there is a PNA or simply ARDS at this time.
4) HTN: Continue Metoprolol to 75 mg po TID and Norvasc 10mg PO daily. BP continues to be elevated in the setting of his agitation, as does heartrate. This may be multifactorial in nature, with dehydration, fevers, and pain contributing. It is notable that his agitation did not improve significantly to prn ativan and pain medications.
5) Anemia: Hct stable
6) s/p cardiac transplant: Cards, Dr. , following. Pt's Tacrolimus' levels sporadic. Continually readjusting dosage. Today's level pendir Per pharmacy goal 8-10. Continue to hold TF while administering Tacrolimus.
7) Seizure: Will continue Keppra.
8) DVT prophylaxis: Pt with hx of HITs, but not documented as occurring this hospitalization. Remains on SCDs and Teds.
9) Volume Status: Increased fluid intake over past 24 hours. Net +527ml, but BUN slightly worse at 33 from 26, suggesting some degree of dehydration; continue to hydrate liberally.
10) FEN: TF at goal rate. GI willing to place PEG when appropriate for long term care. Awaiting for improvement in WBC and low-grade fever before placiing PEG.
11) Dispo: Needs continued ICU care - not ready for transfer to SNF. Need to readdress current status with family.
12) CODE: FULL
By type
By status
By date
Performed by
By encounter
Action
Performed By
Performed Date
Action Status
Comment
Proxy Personnel
Requested By
Requested Date
Perform
Sign
Verify
Sign
Modify
5/23/2006 5:14 PM
Completed
5/23/2006 8:37 PM
Completed
5/23/2006 8:37 PM
Completed
5/23/2006 5:14 PM

Specific Aims

Specific Aim 1

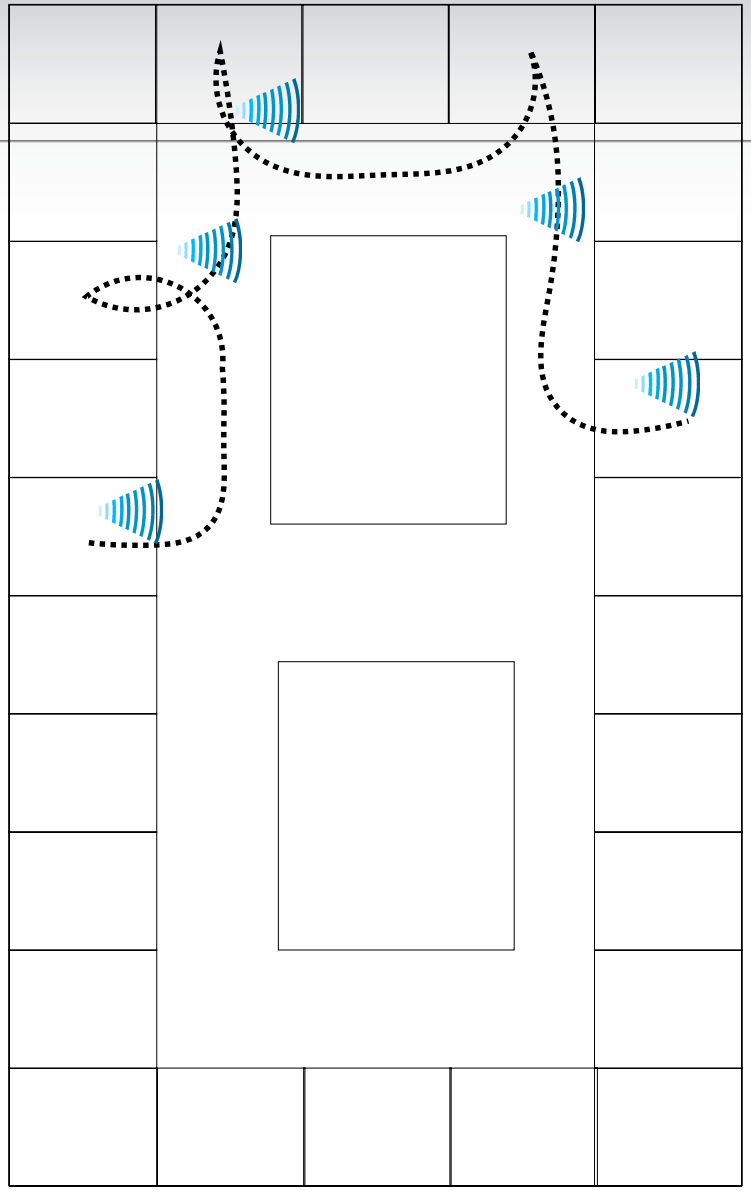
To refine and implement a new **voice**-generated enhanced electronic note system (VGEENS), integrating voice recognition and transcription with natural language processing and links to the electronic medical record (EMR) to improve note creation efficiency and note accuracy.

Specific Aim 2

To evaluate VGEENS using a **randomized trial** with 30 internal medicine physicians in each arm to assess electronic note creation efficiency, note accuracy, and user satisfaction. Intervention physicians will use VGEENS, while control physicians will continue with note creation as they normally would.

Specific Aim 1

Notes will be ‘dictated’ at the bedside or immediately after leaving it.

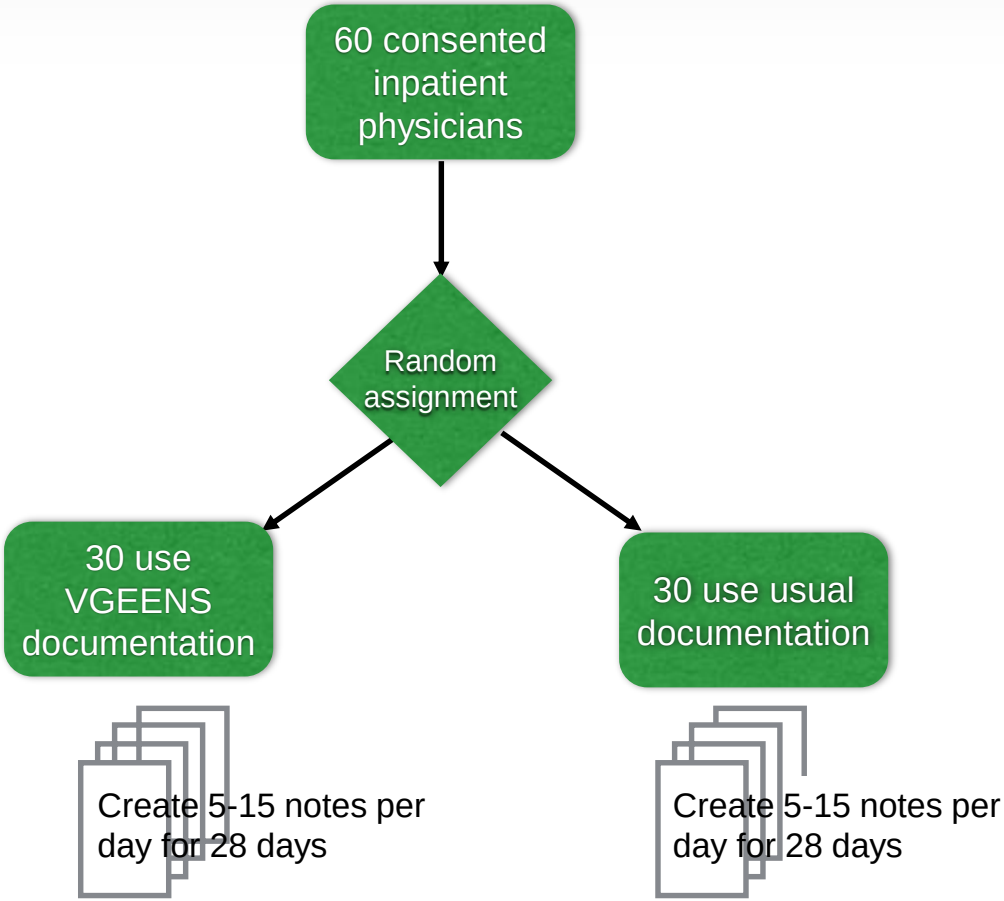


Ways to Write Inpatient Progress Notes

	Fast	Low cost	Incorporates EMR data	Encoded data	Well suited to rounding workflow
VGEENS	✓	✓	✓	✓	✓
Keyboard		✓	✓	✓	
Dictation	✓				✓
Voice recognition		✓			

Randomized Controlled Trial (Planned)

Specific Aim 2



1. Minutes between when patient seen on rounds and note signed in EMR
2. Note quality, measured by Physicians Documentation Quality Instrument
3. Satisfaction of physician users

Physicians Documentation Quality Instrument

Attribute	Score	Description of Ideal Note
1. Up-to-date	Not at all 1 2 3 4 Extremely 5	The note contains the most recent test results and recommendations.
2. Accurate	Not at all 1 2 3 4 Extremely 5	The note is true. It is free of incorrect information.
3. Thorough	Not at all 1 2 3 4 Extremely 5	The note is complete and documents all of the issues of importance to the patient.
4. Useful	Not at all 1 2 3 4 Extremely 5	The note is extremely relevant, providing valuable information and/or analysis.
5. Organized	Not at all 1 2 3 4 Extremely 5	The note is well-formed and structured in a way that helps the reader understand the patient's clinical course.
6. Comprehensible	Not at all 1 2 3 4 Extremely 5	The note is clear, without ambiguity or sections that are difficult to understand.
7. Succinct	Not at all 1 2 3 4 Extremely 5	The note is brief, to the point, and without redundancy.
8. Synthesized	Not at all 1 2 3 4 Extremely 5	The note reflects the author's understanding of the patient's status and ability to develop a plan of care.
9. Internally Consistent	Not at all 1 2 3 4 Extremely 5	No part of the note ignores or contradicts any other part.
Total Score:		

Results

VGEENS System Successfully Developed

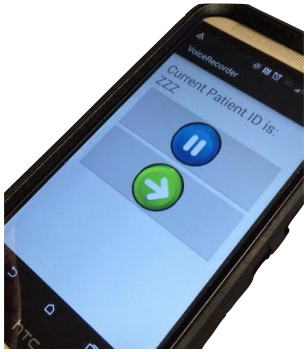
Specific Aim 1

- Used with commercial EHR
- Notes available in EHR Inbox within 5 minutes
- Secure
- Enhanced with text processing later in trial
- Downtime uncommon

vgeens

Voice Generated Electronically Enhanced Notes

Specific Aim 1



Voice dictation files will then be securely transmitted to servers…

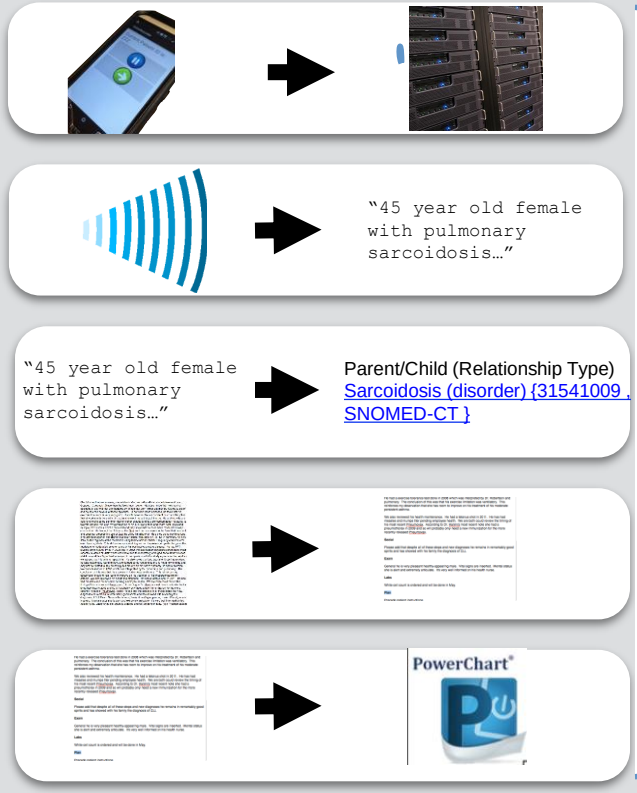
Specific Aim 1

Physician time

Record note (~5 minutes)

Edit and sign note (~3 min)

Computer time

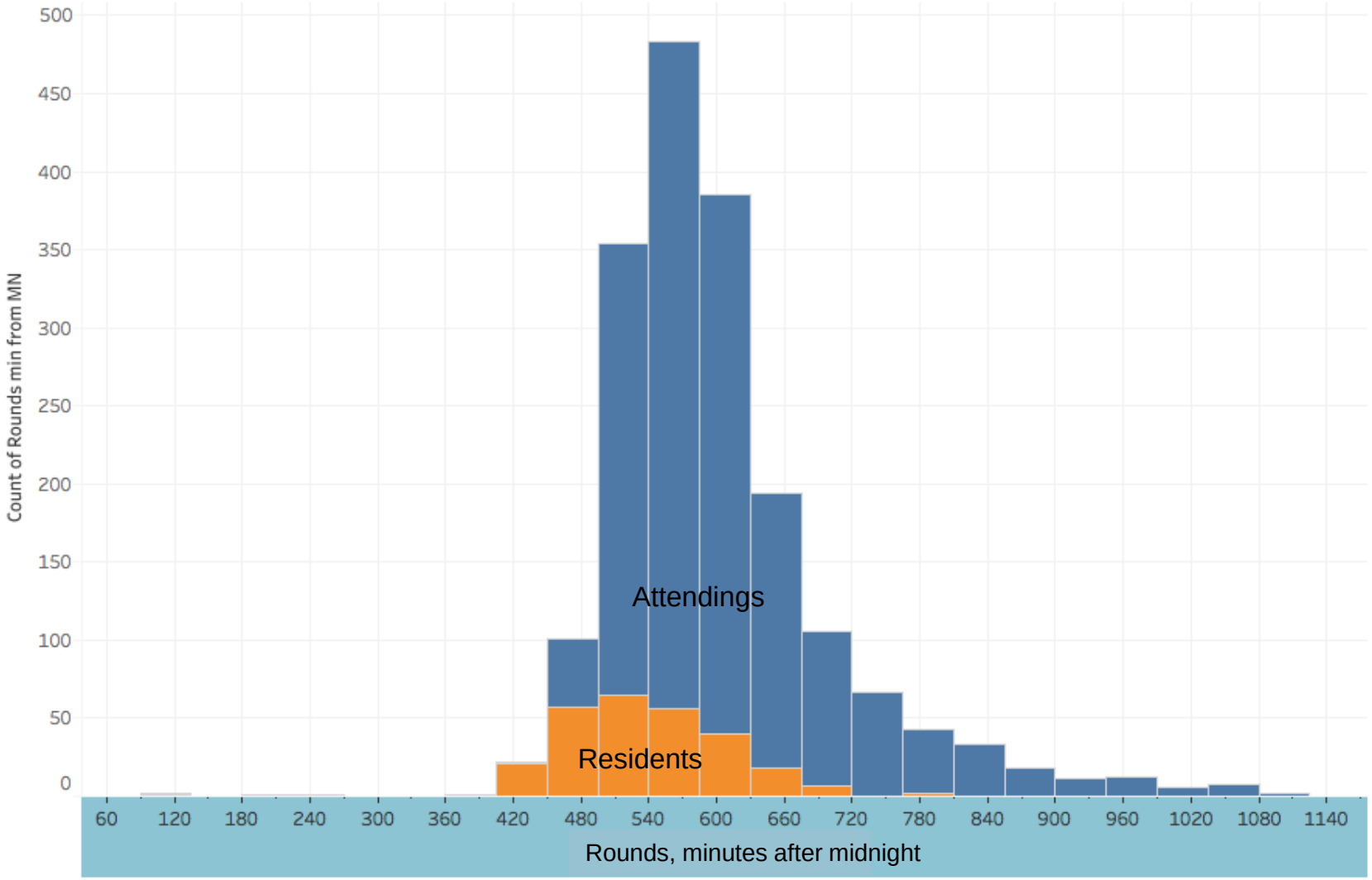


Summary of Subjects, Outcome Data, and Results

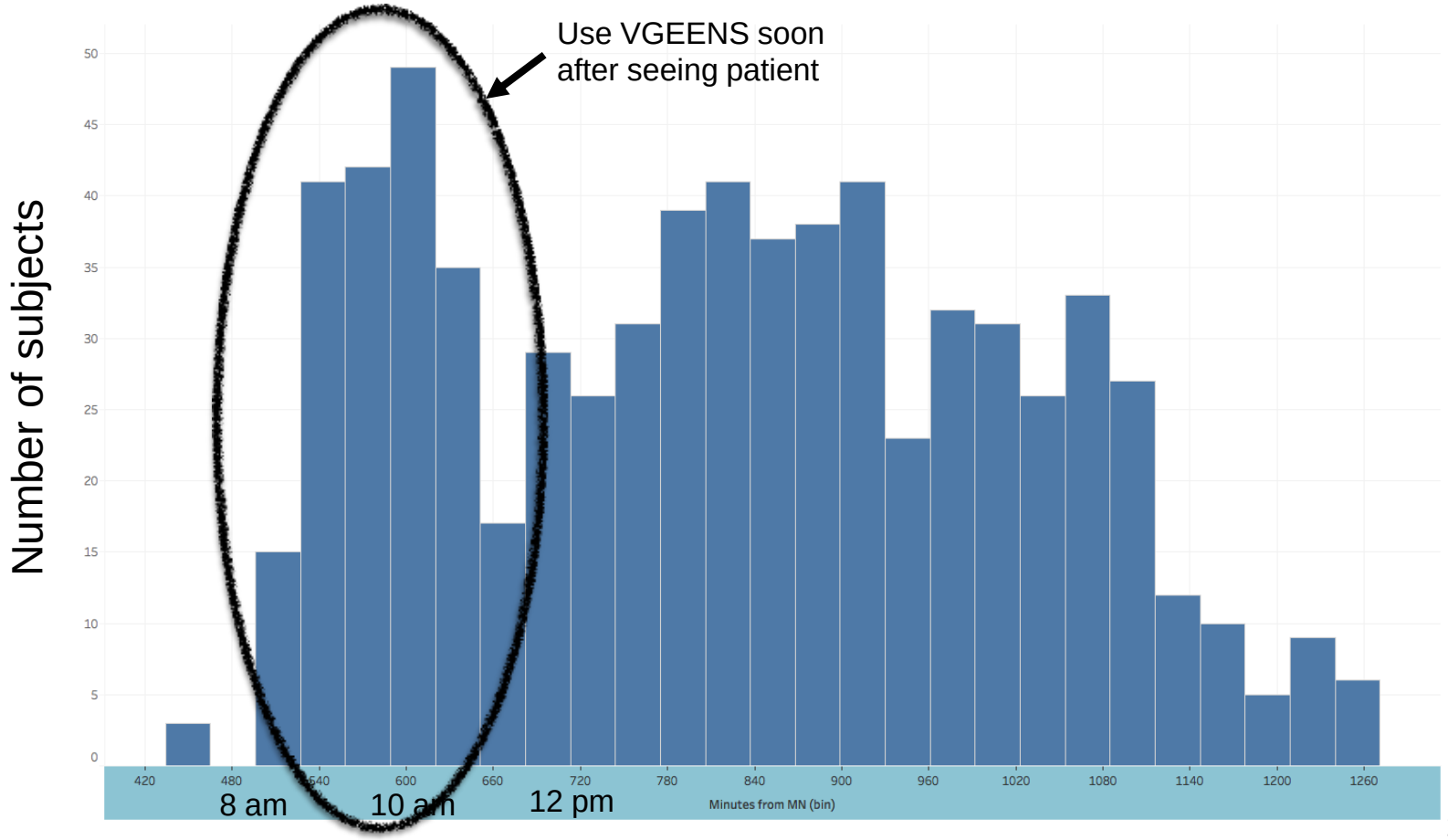
Specific Aim 2

	Intervention	Control
SUBJECTS		
Consented	24	25
Wrote ≥ 1 note (remainder results from these)	13	18
OUTCOME DATA		
Recorded rounding time (%)	99	99
Notes written	709	1143
Satisfaction survey response (%)	100	100
OUTCOME RESULTS		
Timing: Note available – Rounding time (minutes)	227	190
Satisfied: Highly or moderately (%)	40	50
Dissatisfied: Moderately or dissatisfied (%)	40	6
Note quality: Pending		

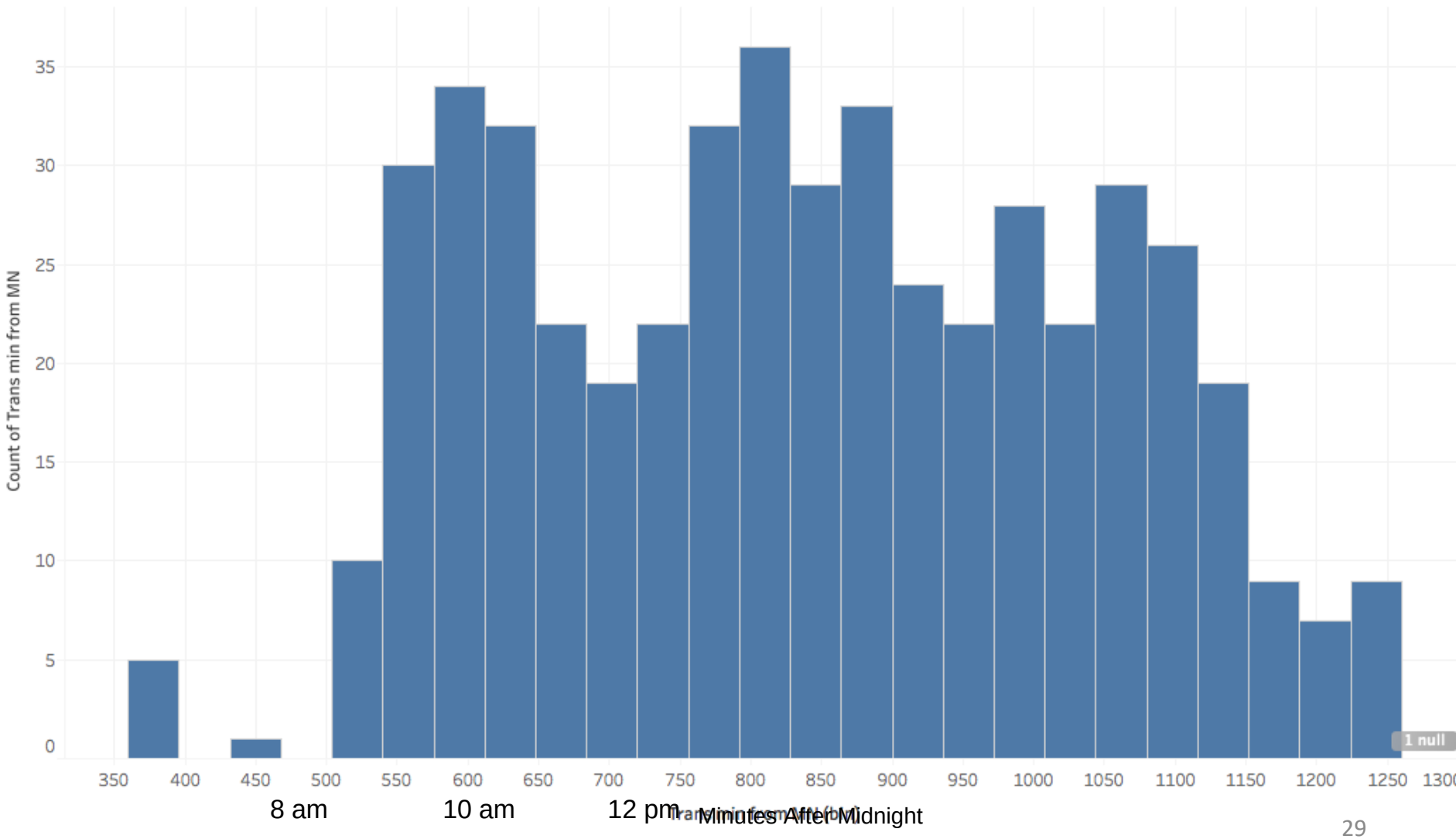
Number of Minutes From Midnight When Patients Seen on Rounds



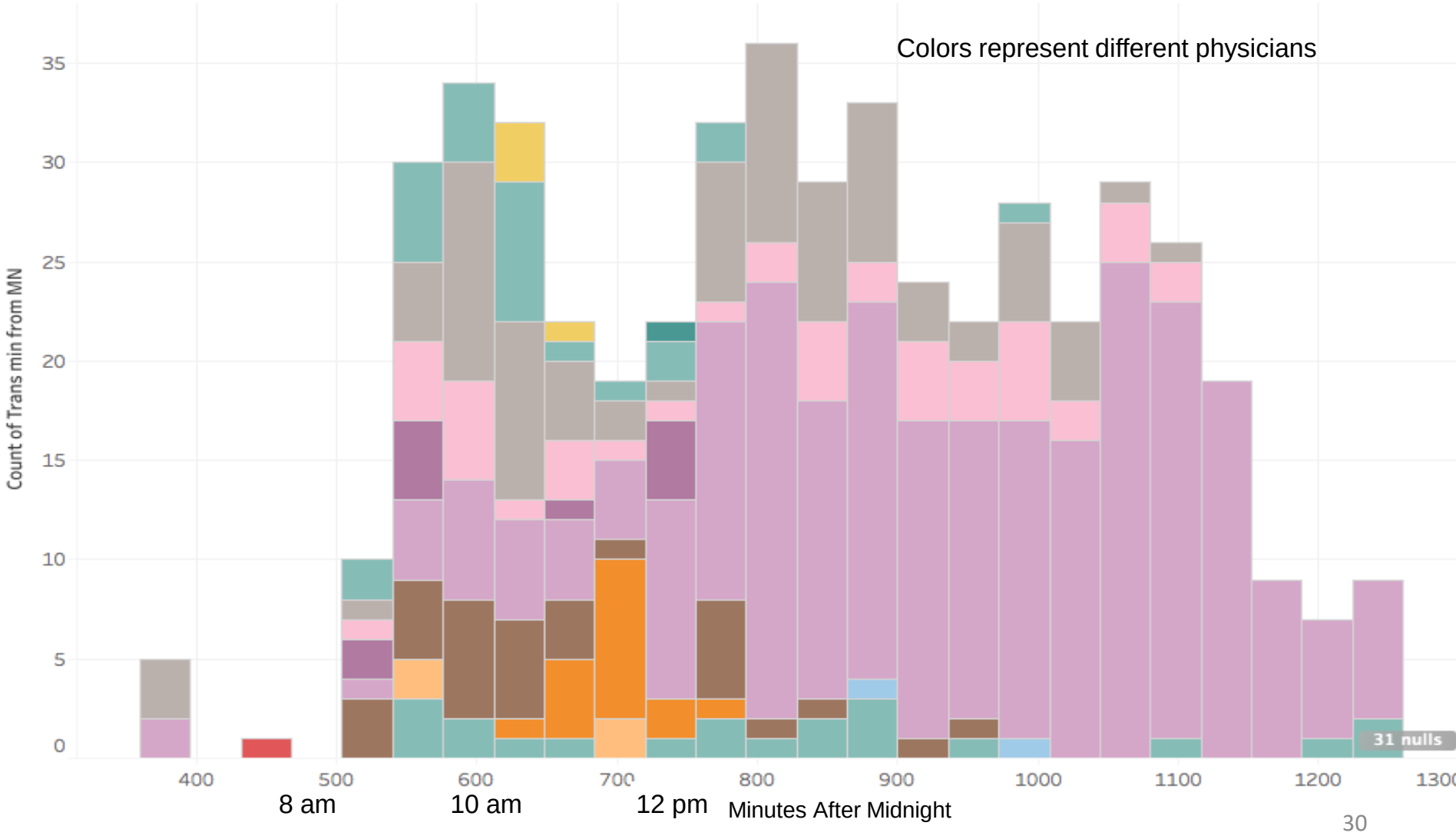
INTERVENTION: Number of Minutes From Midnight When VGEENS Used



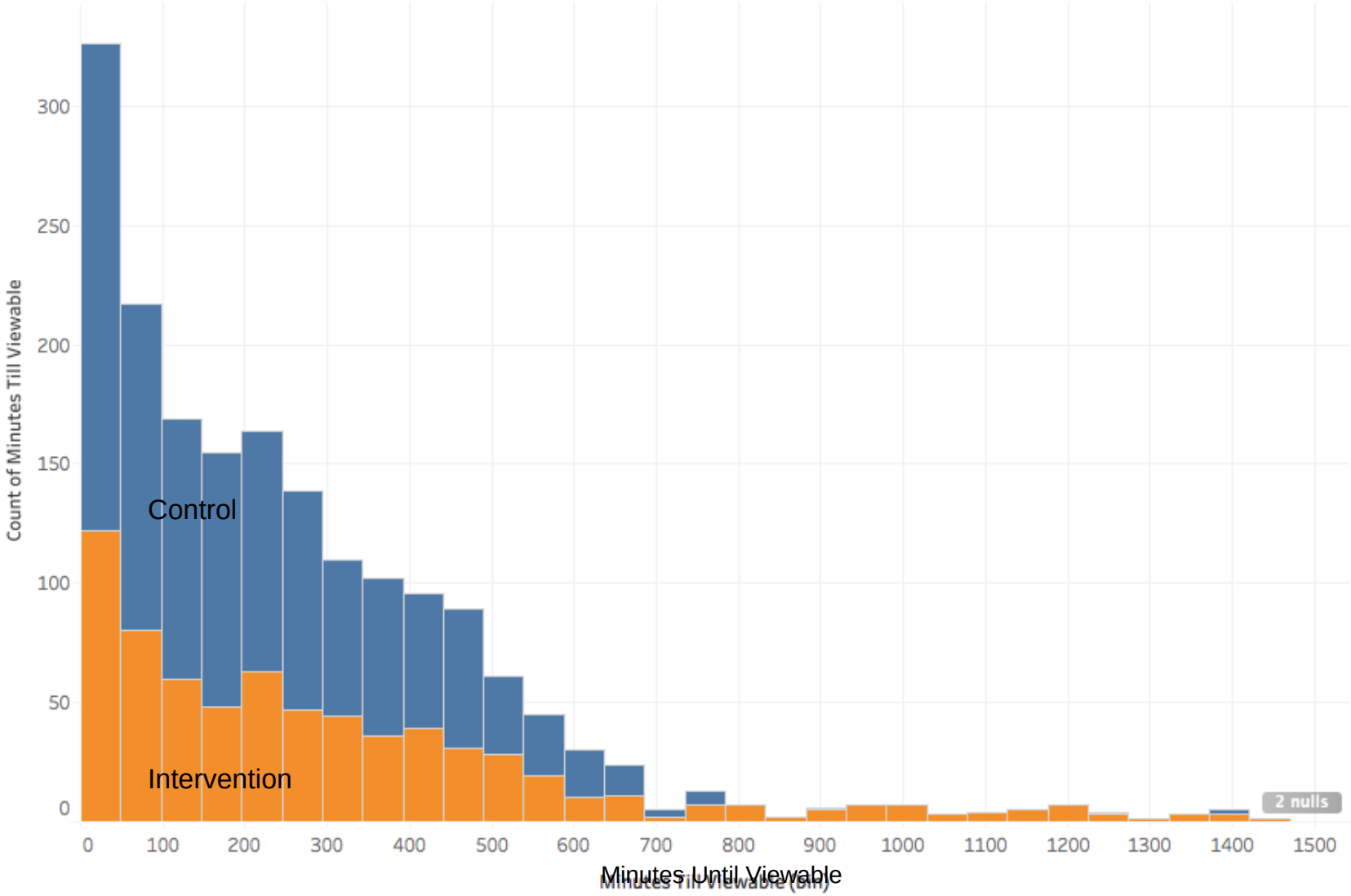
Number of Minutes From Midnight When VGEENS Notes Available



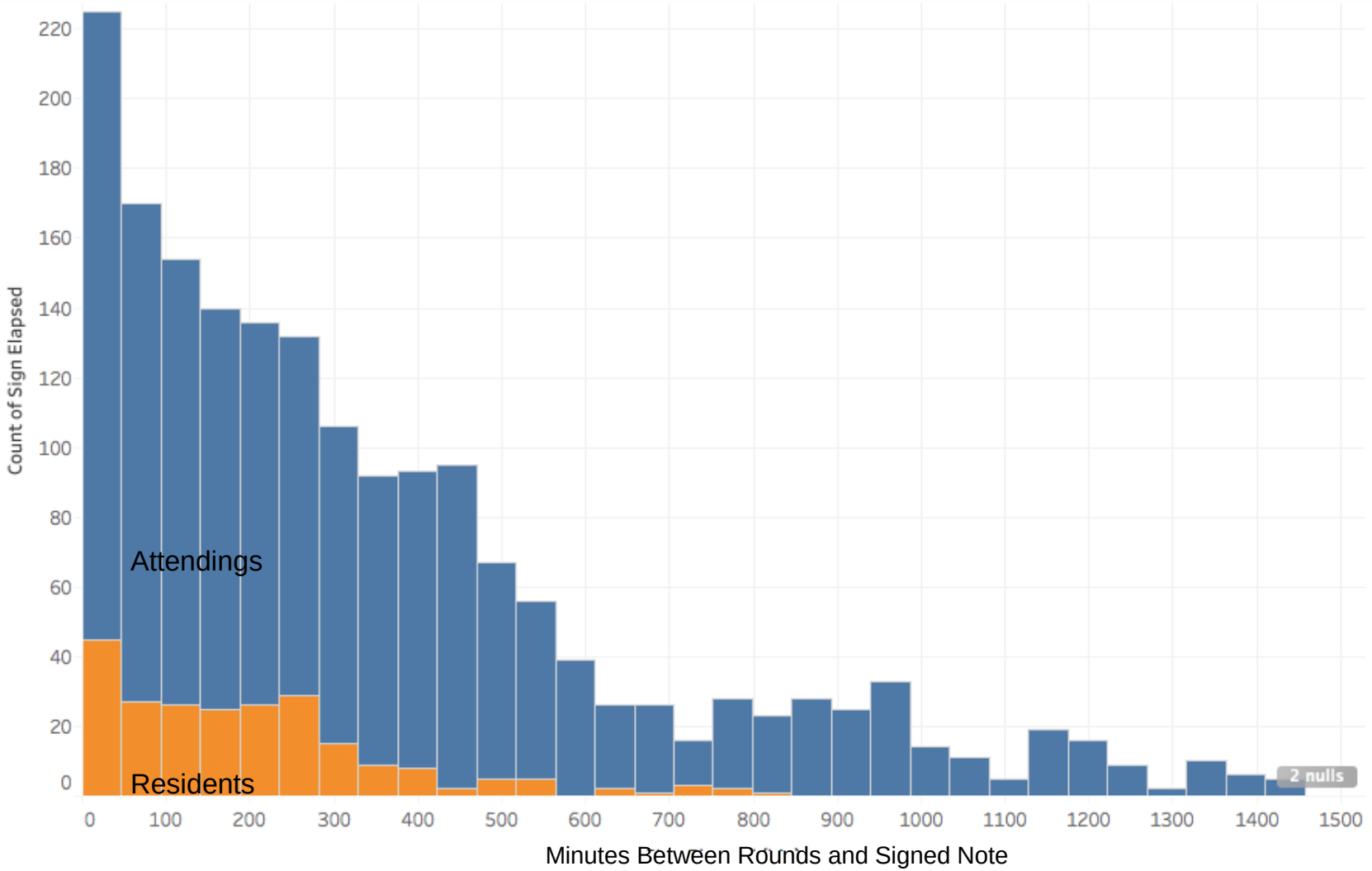
Minutes After Midnight Notes Transcribed, by Author



Time Until Note Viewable Intention to Treat

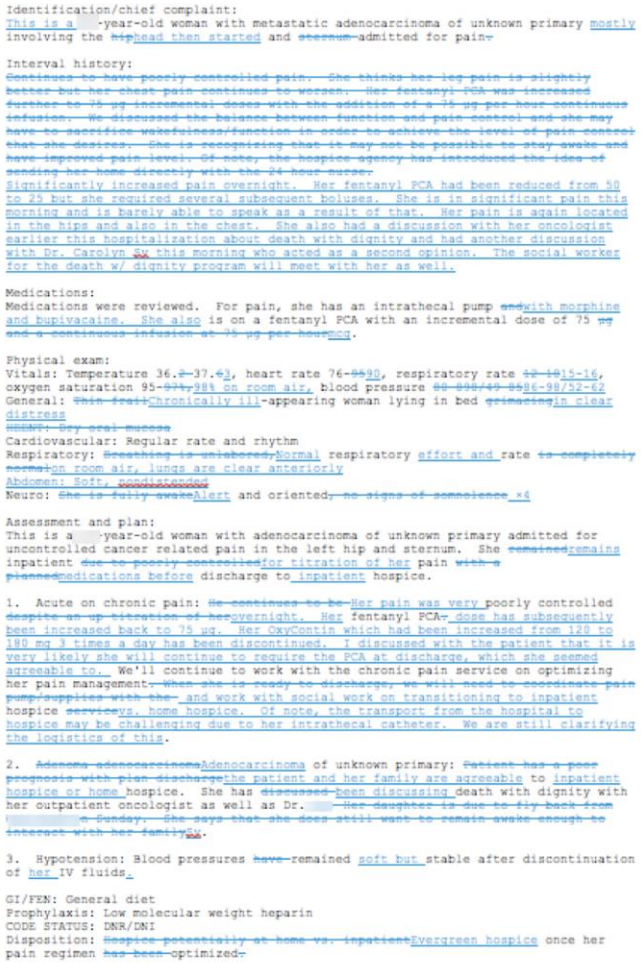


Time Between Rounds and Signed Note (Control)



Comparing VGEENS and Control Notes With Note Written Previous Day on Same Patient

VGEENS



Identification/chief complaint:
This is a -year-old woman with metastatic adenocarcinoma of unknown primary mostly involving the ~~hip~~head then started and ~~sternum~~ admitted for pain~~.~~

Interval history:
~~Continues to have poorly controlled pain. She thinks her leg pain is slightly better but her chest pain continues to worsen. Her fentanyl PCA was increased further to 75 µg incremental doses with the addition of a 75 µg per hour continuous infusion. We discussed the balance between function and pain control and she may have to sacrifice wakefulness/function in order to achieve the level of pain control that she desires. She is recognizing that it may not be possible to stay awake and have improved pain level. Of note, the hospice agency has introduced the idea of sending her home directly with the 24 hour nurse.~~
~~Significantly increased pain overnight. Her fentanyl PCA had been reduced from 50 to 25 but she required several subsequent boluses. She is in significant pain this morning and is barely able to speak as a result of that. Her pain is again located in the hips and also in the chest. She also had a discussion with her oncologist earlier this hospitalization about death with dignity and had another discussion with Dr. Carolyn ~~Sx~~ this morning who acted as a second opinion. The social worker for the death w/ dignity program will meet with her as well.~~

Medications:
Medications were reviewed. For pain, she has an intrathecal pump ~~and~~with morphine and bupivacaine. She also is on a fentanyl PCA with an incremental dose of 75 ~~ug~~ and a continuous infusion at 75 µg per hour~~mcg~~.

Physical exam:
Vitals: Temperature 36.~~2~~-37.~~3~~, heart rate 76-~~99~~90, respiratory rate ~~12-18~~15-16, oxygen saturation 95-~~97%,98%~~ on room air, blood pressure ~~88-098/49-86~~96-98/52-62
General: ~~Thin frail~~Chronically ill-appearing woman lying in bed ~~grimacing~~in clear distress
~~HEENT: Dry oral mucosa~~
Cardiovascular: Regular rate and rhythm
Respiratory: ~~Breathing is unlabored,~~Normal respiratory effort and rate ~~is completely normal~~on room air, lungs are clear anteriorly
Abdomen: Soft, ~~nondistended~~
Neuro: ~~She is fully awake~~Alert and oriented~~, no signs of somnolence~~ ×4

Assessment and plan:
This is a -year-old woman with adenocarcinoma of unknown primary admitted for uncontrolled cancer related pain in the left hip and sternum. She ~~remained~~remains inpatient ~~due to poorly controlled~~for titration of her pain ~~with a planned~~medications before discharge to inpatient hospice.

1. Acute on chronic pain: ~~He continues to be~~Her pain was very poorly controlled ~~despite an up titration of her~~overnight. Her fentanyl PCA~~,~~ dose has subsequently been increased back to 75 ug. Her OxyContin which had been increased from 120 to 180 mg 3 times a day has been discontinued. I discussed with the patient that it is very likely she will continue to require the PCA at discharge, which she seemed agreeable to. We'll continue to work with the chronic pain service on optimizing her pain management~~. When she is ready to discharge, we will need to coordinate pain pump/supplies with the~~ and work with social work on transitioning to inpatient hospice ~~service~~vs. home hospice. Of note, the transport from the hospital to hospice may be challenging due to her intrathecal catheter. We are still clarifying the logistics of this.

2. ~~Adenoma adenocarcinoma~~Adenocarcinoma of unknown primary: ~~Patient has a poor prognosis with plan discharge~~the patient and her family are agreeable to inpatient hospice or home hospice. She has ~~discussed~~ been discussing death with dignity with her outpatient oncologist as well as Dr. ~~ Her daughter is due to fly back from ~~ on Sunday. She says that she does still want to remain awake enough to interact with her family~~Sx~~.

3. Hypotension: Blood pressures ~~have~~ remained ~~soft but~~ stable after discontinuation of her IV fluids.

GI/FEN: General diet
Prophylaxis: Low molecular weight heparin
CODE STATUS: DNR/DNI
Disposition: ~~Hospice potentially at home vs. inpatient~~Evergreen hospice once her pain regimen ~~has been~~ optimized~~.~~

Control

INPATIENT PROGRESS NOTE

HOSPITAL DAY: ~~3~~4

IDENTIFICATION/CHIEF CONCERN: with depression admitted after suicide attempt by ingestion

INTERVAL HISTORY:
~~Extubated and delivered by MHP yesterday~~
~~Pain much better previously today but Fetis happy~~
~~"Even in jail I wasn't this sad"~~
~~No events overnight~~
~~Feeling fine, just wants to go home~~

ALLERGIES:
NKA

SCHEDULED MEDICATIONS:
~~heparin 5,000 units/mL(5000) 40 mg~~ Subcutaneous ~~Q8 Hours~~QHS
magnesium oxide 800 mg PO Daily

PHYSICAL EXAM:
(02/~~13~~/16 11:~~30~~0) T 36.~~4~~5 HR ~~88~~ RR ~~16~~ BP ~~108/62~~107/68 MAP 76 O2 Sat 98% on RA
~~(02/15/16 08:00 - 02/16/16 07:59)~~
In: ~~3000~~; Out: 0; Net: ~~3000~~ *(last 24 hours)*

GEN: young woman sitting up in bed, awake and alert. Smiling and engaged with ~~family, initially engaged~~boyfriend, polite cooperative with me ~~but when I talked a little more about her situation showed eye contact and verbal responses~~
HEENT: PERRL, sclera anicteric, OP clear w/MMM
CV: RRR no ~~murmurs~~
~~Resp~~ CTA ~~bl~~ with normal respiratory effort
~~Abd~~ +BT soft NT ND
~~Extrem~~ warm, no edema, locked restraint L ~~arm~~ R ankle
Skin: warm and dry
Neuro/psych: minimal significant engagement, cooperative with exam, brighter affect, MAE

LABS: *(Most recent results in ~~24 hour~~ range.)*

RESULTS FROM TODAY

~~WBC~~	~~Hgb~~	~~Hct~~	~~Plt~~	~~Na~~	~~K~~	~~Cl~~	~~CO2~~	~~BUN~~	~~Cr~~	~~Glc~~	~~Ca~~
~~---~~	~~---~~	~~---~~	~~---~~	137	3.8	102	26	14	0.61	~~---~~	~~8.8~~

~~Mg~~	~~---~~	~~Phos~~	~~---~~
~~1.8~~	~~---~~	~~3.5~~	~~---~~

RESULTS FROM YESTERDAY ~~02/16/16 05:58~~

~~WBC~~	~~Hgb~~	~~Hct~~	~~Plt~~	~~Na~~	~~K~~	~~Cl~~	~~CO2~~	~~BUN~~	~~Cr~~	~~Glc~~	~~Ca~~
~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	9.8

~~Mg~~	1.7	~~Phos~~	~~---~~
~~---~~	~~---~~	~~---~~	~~---~~

RESULTS FROM TODAY

~~PT~~	~~---~~	~~PTT~~	~~---~~	~~INR~~	~~---~~
~~---~~	~~---~~	~~---~~	~~---~~	~~---~~	~~---~~

RESULTS FROM YESTERDAY ~~02/16/16 05:58~~

~~PT~~	10.7	~~PTT~~	23	~~INR~~	1.0

No CBC Results Found.
~~salicylate level 2/17 < 3~~

ASSESSMENT & PLAN:

with depression/anxiety and h/o prior suicide attempts admitted to the ICU 2/14 after attempting suicide by ingesting diphenhydramine, ASA, and ~~tylenol~~. MHP detained patient, medically cleared for psych when bed available.

Suicide attempt: She had called her boyfriend and reported taking 48 tablets of ~~benadryl~~, 44 tabs of ~~excedrin~~ PM (~~benadryl~~, ASA, ~~tylenol~~). Unclear trigger. Declining to engage with psych for ~~eval~~ so MHP called. Boyfriend is at bedside, appears very supportive. Toddler daughter is with patient's mother, came in for visit ~~today~~ 2/17. Patient was very happy and engaged with her baby but still not willing to engage in anything but very superficial conversations with staff.
- await psych bed
- ~~court date 2/19~~
- continue locked restraints while in ICU, plan for sitter ~~will transitioned to~~ floor
- psych following

Overdose: Poison Control Center contacted on admission, continuing to follow.
~~Anticholinergic toxicity~~: diphenhydramine. Had AMS, dilated pupils, agitation, warm/dry skin. QRS has not been widened. ~~QTc~~ had a max of 585 on 2/14, down to 423 today. HR was as high as 150s, now stable in 70s. Rec'd ~~physostigmine~~ 1mg x 2 doses since admission for clinical signs of toxicity (tachycardia and AMS). EKG now normal. ~~Repeat mag slightly low~~hypomagnesemia and hypokalemia resolved with repletion. Electrolytes should remain normal with PO intake.
~~- mag oxide 800mg PO x 1~~
- encourage good PO intake

~~APAP ingestion~~: Level at 2h was 29 -->30-->25 -> 15. NAC loading dose given in ER after 2h level. Given nomogram, no maintenance dose recommended. LFTs have been normal.
- no further monitoring needed

~~Salicylate ingestion~~: levels 10, 10, 9 over first 6h here~~, no risk anymore~~. VBG pH 7.4. ~~Recheck on 2/17 was < 3.~~
~~- recheck chem now, if mag at our requires enteral feeding otherwise off pump duties~~
~~- ok to resume bowel/feeding per poison control~~

Leukocytosis: 14 on admission, normalized with supportive care. No fever or localizing signs to suggest infection. Likely ~~demargination~~ with acute ingestion and vomiting.

Fluids/electrolytes/nutrition: regular diet
Prophylaxis: ~~Lovenox~~ SQ

What Is Medical Informatics?
(attributed to Homer Warner)

Technology	10%
Medicine	10%
Sociology	80%

Lessons Learned

- Physicians may resist changing established note-writing habits.
- The VGEENS approach can make creating progress notes faster if voice recorded at bedside or soon after.
- Notes created using voice may contain less text carried forward from prior notes and may be more accurate.
- Features popular with physicians: carrying forward plan, ‘checklist’ information, minimizing editing requirements.

Successes

- We developed and deployed a new system, integrated with a commercial EHR, to create inpatient progress notes within 5 minutes.
- If used at the bedside or soon thereafter, notes are available much sooner for others to view.
- Notes created using voice may contain less text carried forward from prior notes.
- We have a method to apply decision support based on progress note content within minutes.

Challenges

- Physicians may resist changing established note-writing habits.
- On average, satisfaction was greater with usual (control) method of note writing, perhaps because popular VGEENS features weren't available until late in the controlled trial.

Lessons From a First-Time Principal Investigator

- Developing a system to be used in producing and conducting an RCT in 2 years is ambitious, but possible.
- People are enthusiastic about joining an interesting project. Diverse disciplines help!
- Pick a problem you know is important and understand deeply.
- Think about the next project from Day 1 (automated editing, NLP tools, measuring note accuracy, etc.)

Our Team

Co-investigators

Andrew White

Meliha Yetisgen

Tom Gallagher

Computing engineer

Andrew Markiel

Research team

Amelia Chappelle

Jennifer Zech

Collaborators

David Alonso

Xinran (Leo) Liu

Ross Lordon

Kevin J. Lybarger

Mari Ostendorf

Trevor Steinbach

Thanks to AHRQ!

This project was supported by grant number R21HS023631 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality or the Department of Health and Human Services.

Contact Information

Thomas Payne

tpayne@u.Washington.edu

U.S. Department of Health and Human Services
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Improving Health IT Safety Through the Use of Natural Language Processing to Improve Accuracy of EHR Documentation

Li Zhou, M.D., Ph.D.
Harvard Medical School
Brigham and Women's Hospital

Document Quality & Patient Safety

- Accurate medical documents are critical for safe patient care and effective inter-provider communication.
- Errors in medical documents can lead to medical errors in patient care, some of which cause injury or even death.
- ~5 million errors per year are tied to wrong medications; 1 in 4 medication errors involves a pair of drugs whose names look alike or sound alike.
  - *Altenol* vs. *Atenolol*
  - *Lyrica* vs. *Lamictal*

http://www.nbcnews.com/id/37386398/ns/health-health_care/t/look-alike-sound-alike-drugs-trigger-dangers/#.WA0LZOUrLIU

Spelling Errors in Clinical Documents

- **Non-word errors** (e.g., *Hum**u**log* for *Hum**a**log*)
  - Free-text entries, typed notes

- **Real-word error**s (i.e., the word is spelled correctly but is contextually wrong, such as *there* for *their*)
  - Speech Recognition (SR) generated text

Background: Non-Word Errors Detection and Correction

- Spelling errors in free-text EHRs

	Notes (n=315)	Free-text allergy entries (n=2626)	Free-text med orders (n=2743)
Error rate	0.5%	4.5%	7.5%
Clinical terms	28.2%	65.5%	78.0%
Real-word errors	3.8%	1.8%	0%

- We developed a spell checker in our MTERMS NLP system

	MTERMS Spell Checker			Aspell Default
	Notes	Allergies	Medications	Notes
Precision	71.1	96.2	90.0	48.9
Recall	81.0	92.7	91.5	82.3
F measure	75.7	94.4	90.8	61.3
Accuracy	78.1	88.2	81.5	58.5

Lai KH, Topaz M, Goss FR, Zhou L. Automated misspelling detection and correction in clinical free-text records. *Journal of Biomedical Informatics*. 2015 Jun;55:188-95. PMID: 25917057

Background: Speech Recognition and Real-Word Errors

- Previous studies were limited in scope and sample size
- Error rates by word range from 1.5%-15%
- High error rate by report (23%) were found in radiology reports
- 76% of radiologists believed error rate by report < 10%

Author	Doc Type	Sample Size	Error Rate
Devine* 2000	A discharge summary and a progress note	12 physicians	IBM software: 7.0% to 9.1% Dragon: 14.1% to 15.2%
Kanal* 2001	72 radiology reports	6 participants	IBM MedSpeaker: 10.3% Significant errors: 7.8%
Zick* 2001	47 ED charts	2 physicians	Dragon NaturallySpeaking: 1.5% Errors/chart: 2.5
Quint 2008	265 radiology reports	-	22% of reports contained significant errors
Basma 2011	615 radiology reports	-	23% of SR reports contained major errors 4% in conventional dictation transcriptions

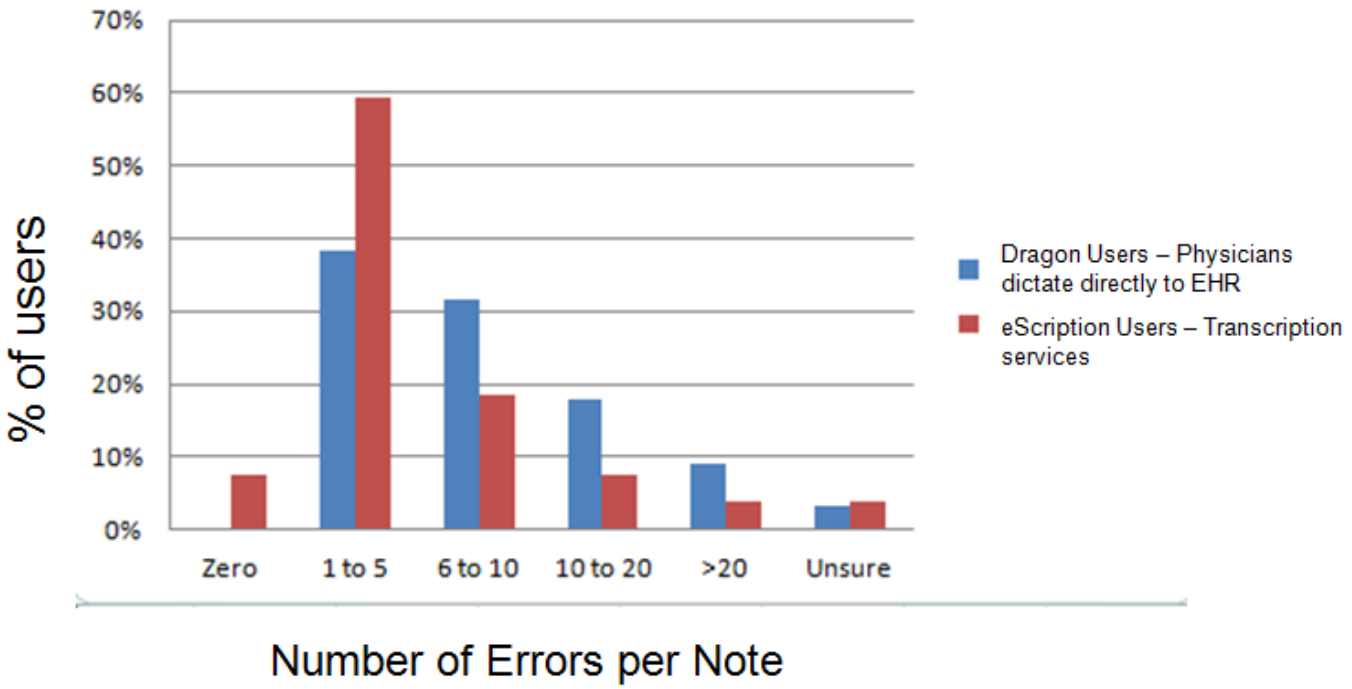
*** Error Rates:** total number of errors divided by the total number of words in the report.

Background: Our SR Error Study in ED

- Retrospectively analyzed 100 emergency physician notes during Dec 2012.
  - Generated via a front-end SR system (Dragon® Medical 10.0)
  - Further edited and signed by the physicians
- 71% of notes contained errors; 1.3 errors per note; 9 errors per 1000 words.
- 15% contained one or more clinically significant errors.
- Physicians signed their notes with known errors, indicating proofreading the entire medical note to search for errors is time consuming.

Goss F, Zhou L, Weiner S. Incidence of Speech Recognition Errors in the Emergency Department. *Int J Med Inform*. 2016 Sep; 93:70-3.. Epub 2016 May 26. PMID: 27435949.

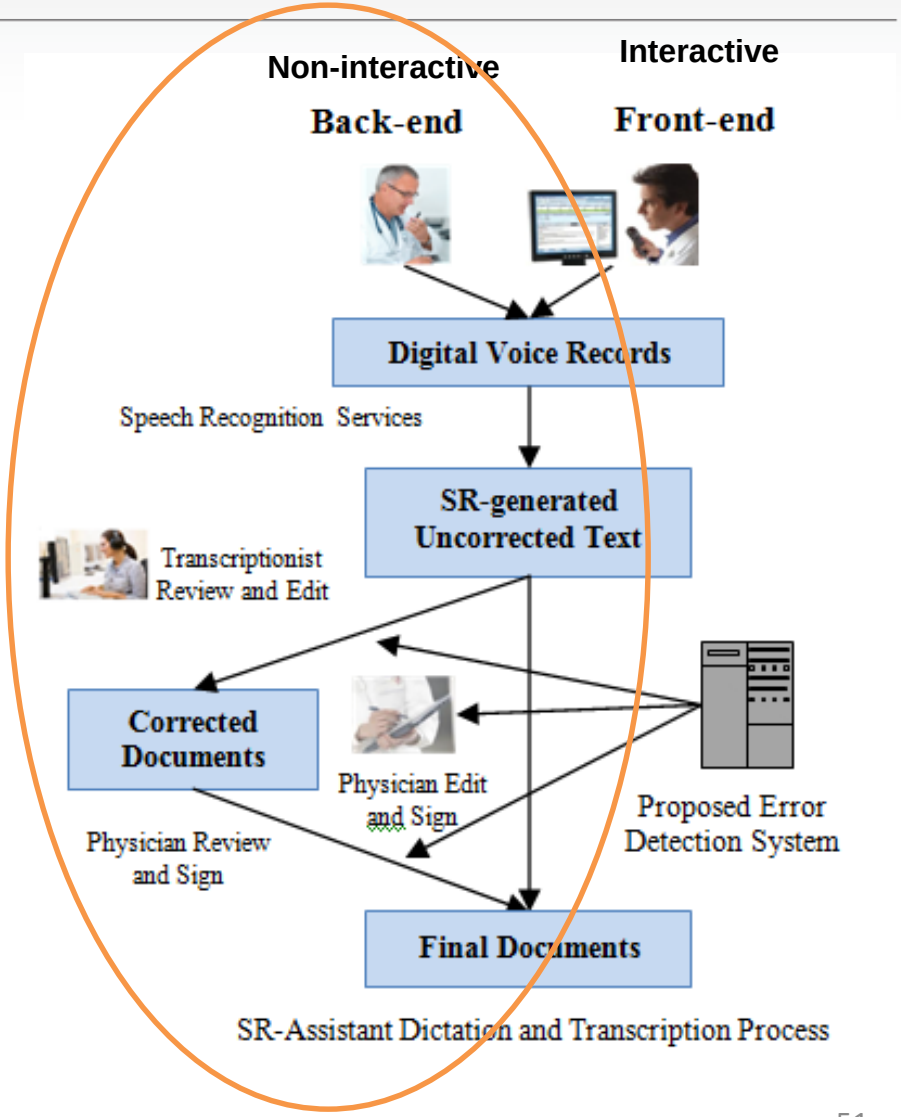
Background: Our Clinician Survey



Surveyed 114 Dragon Users and 50 eScription Users at Brigham and Women's Hospital, Boston

SR Error Analysis and Detection
(Ongoing Study Funded by AHRQ)

- **Aim 1:** Conduct error analysis to estimate the prevalence and severity of SR errors.
- **Aim 2:** Develop NLP methods for automated error detection.
- This presentation reports our errors analysis in **back-end** SR generated documents at different processing stages.



Methods

- Stratified random sample of 169 dictated notes using transcription services (back-end SR)
  - 79 from Brigham and Women's Hospital (24 operative notes and 55 office notes)
  - 40 discharge summaries from North Shore Medical Center, Boston
  - 50 from the University of Colorado Hospital (35 discharge summaries and 15 operative notes)
- Four processing stages
  - Original audio file dictated by the provider (**AO** note)
  - Note generated by SR engine of the vendor transcription service (**SR** note)
  - Note edited by a professional medical transcriptionist (**MT** note)
  - Final note reviewed and signed by a clinician (**SN** note)
- Three-level annotation schema
  - General error types
  - Semantic error types
  - Clinically significant errors
- Manual review to create gold standard

Measures

- **Length of time** to dictate a note
- **Turnaround time** for each note version
- **Differences** in the SR note, MT note, and SN note **from the gold standard**
- **Error rate**: number of errors divided by the number of words
- **Percentage of each error type** by overall errors
- Percentage of notes with at least one **clinically significant error**
- **Repeated** these analyses for SR, MT, and SN notes; for each note type; and across all notes

General Error Types

Type		Description	Example
Insertion		One or more words was added to the transcription	AO: There is distal biliary obstruction observed SR: There is **no** distal biliary obstruction observed
Deletion		One or more words was deleted from the transcription	AO: **CHADS2 VASC** score **4** SR: score
Substitution	Enunciation	An error due to a mispronunciation or failure to enunciate on the part of the speaker	AO: **to find a** homeopathic provider SR: **defined** homeopathic provider
	Suffix error	The root word is correct, but there is an incorrect, added, or omitted suffix	AO: mental status **worsened** SR: mental status **worsens**
	Dictionary error	An error due to the target word not being present in the SR system's dictionary	AO: driving a **Camry** and hit another car SR: driving an **Academy** and hit another car
	Spelling error	The ***transcriptionist*** made a spelling error when editing the output of the SR system	AO: we counseled **him** on risk of infection MT: we counseled **hom** on risk of infection
	Homonym error	One word has been substituted for another identically pronounced word	AO: **serial** high resolution anoscopy SR: **cereal** high resolution anoscopy
	Nonsense error	A substitution that is so far off that it cannot be determined which (if any) category it falls under	AO: **follow up in 3 to 5 days** SR: **neck veins are evaluated**
	Prefix Error	The root word is correct, but there is an incorrect, added, or omitted prefix	AO: **Inadequate** evaluation to exclude neoplasia SR: **Adequate** evaluation to exclude neoplasia
Number error		Any error involving a number, whether it is written as a digit ("3") or as a word ("three")	AO: the patient is a **17**-year-old female SR: the patient is a **70**-year-old female
Punctuation Error		A period, comma, or other punctuation mark was added where it should not have been	AO: at discharge she had no flank tenderness SR: at discharge**.** She had no flank tenderness

Semantic Error Types

Type	Description	Example
General English	Any English words that do not fit into the categories below	AO: which she would otherwise **forget** SR: which she would otherwise **for gas that**
Stop Word	Common English words (we are using the list defined at http://ranks.nl/stopwords)	AO: intermittent pain **under** the right breast SR: intermittent pain **in** the right breast
Medication	Medication names and dose information	AO: initiated on **Lamotrigine** therapy SR: initiated on **layman will try** therapy
Diagnosis	Any words that are part of a specific medical diagnosis	AO: **Dengue** SR: **DKA**
Lab	Includes lab test names and lab test results	AO: TSH of **26.7** SR: TSH of **22nd 6.7**
Imaging Test	Imaging exam names/types and exam results	AO: nonobstructive on **CT** imaging SR: nonobstructive on imaging
Procedure	Procedure names and descriptions	AO: CA ligament was released **on the leading** edge SR: CA ligament was released **operating** edge
Physical exam	Any information directly related to the physical exam (ht/wt, HR, BP, etc.) and any associated values	AO: **T 36.7 degrees** SR: **T3-T7 disease**
Patient/ provider info.	Any words involving patient/provider metadata, such as the patient's name, doctors' names, patient MRN, etc.	AO: SURGEON: **[surgeon's actual name]** SR: SURGEON: **Stathis stairs**
Date	Any dates, including those that are written with words (January 1, 2017) or written with numbers (01/01/2017)	AO: 10/10/**2016** SR: 10/10/**2000**
Symptom	Any symptom or description of symptoms	AO: very **mild** althralgias SR: very **also** arthralgias
???	When ??__?? (or similar) is left in the note, or when something is completely nonsense	AO: no foreign **material** was identified MT: no foreign **??__??** was identified

Preliminary Results

Our study is ongoing; final results may differ from those presented here.

Results

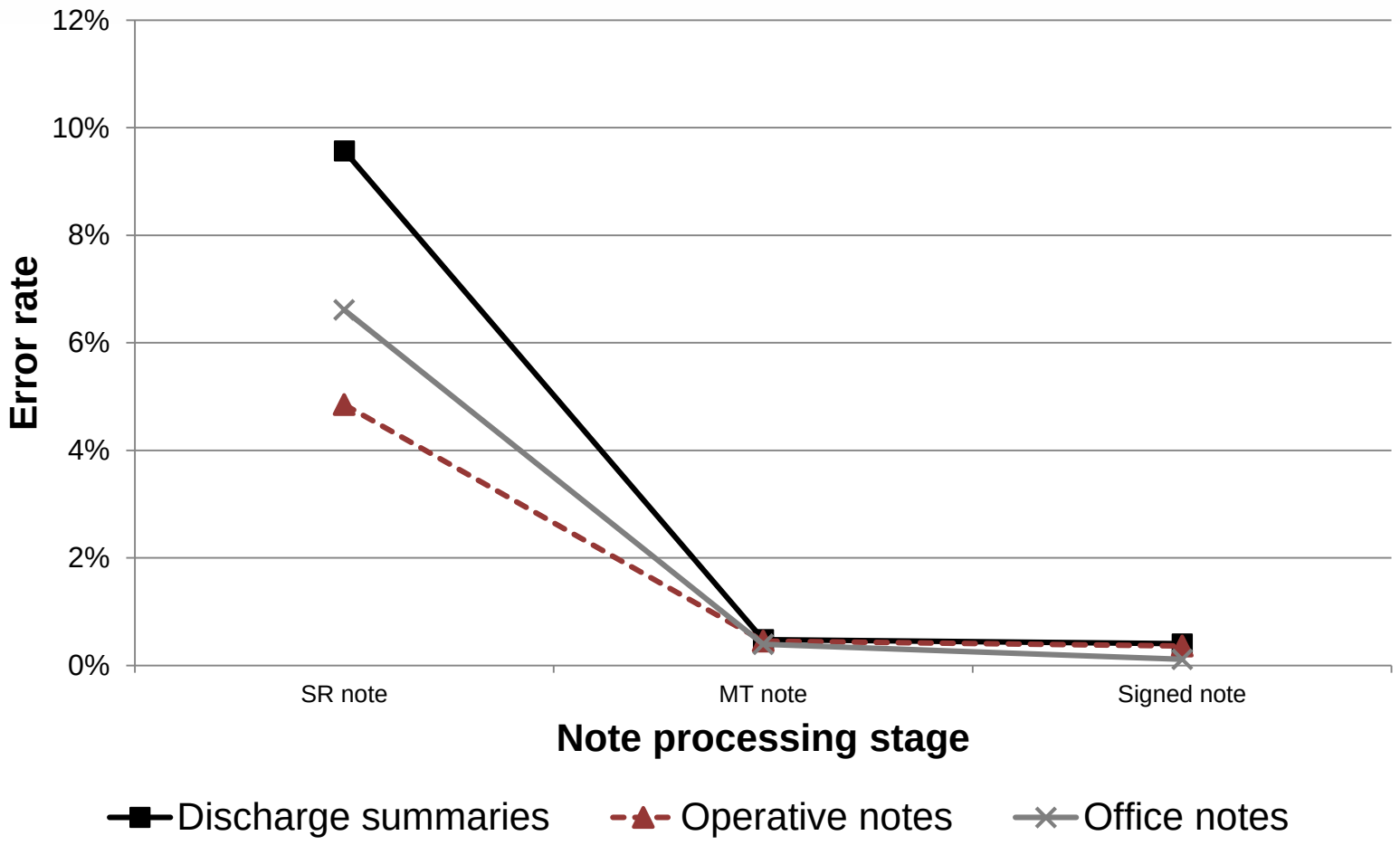
	Mean	Median	Minimum	Maximum
Length of notes (words)	558	524	102	1230
Dictation time (minutes)	5	4.5	0.4	31.5
Turnaround time – Time between completion of dictation and upload to EHR system (hours)	3.5	1	2 minutes	38.8
Clinician review time - Time between upload to EHR system and clinician signing of note (days)	4.2	1	0	42

Overall Error Rates and General Types

		Total Errors n (%)[1]	Errors – General Types n (%)[2]		
			Deletion	Insertion	Enunciation
Discharge Summaries (75)	SR	3892 (9.6)	**1395 (35.8)**	1031 (26.5)	655 (16.8)
	MT	195 (0.5)	87 (44.6)	36 (18.5)	35 (18.0)
	SN	163 (0.4)	74 (45.4)	29 (17.8)	29 (17.8)
Office Notes (55)	SR	1588 (6.6)	**539 (33.9)**	306 (19.3)	431 (27.1)
	MT	96 (0.4)	29 (30.2)	13 (13.5)	36 (37.5)
	SN	32 (0.1)	6 (18.8)	7 (21.9)	12 (37.5)
Operative Notes (39)	SR	1233 (4.8)	376 (30.7)	**401 (32.8)**	167 (13.7)
	MT	120 (0.5)	47 (39.2)	29 (24.2)	18 (15.0)
	SN	96 (0.4)	42 (43.8)	25 (26.0)	15 (15.6)
All Notes (169)	SR	**6703 (7.5)**	**2310 (34.5)**	1738 (25.9)	1253 (18.7)
	MT	**411 (0.6)**	**163 (39.7)**	78 (19.0)	89 (21.7)
	SN	**291 (0.3)**	**122 (41.9)**	61 (21.0)	56 (19.2)

[1] n = number of errors; % = total number of errors divided by the total number of words in the notes.
[2] n = number of errors; % = number of errors of a specific type divided by the total number of errors.

Error Rates across All Note Types and Stages

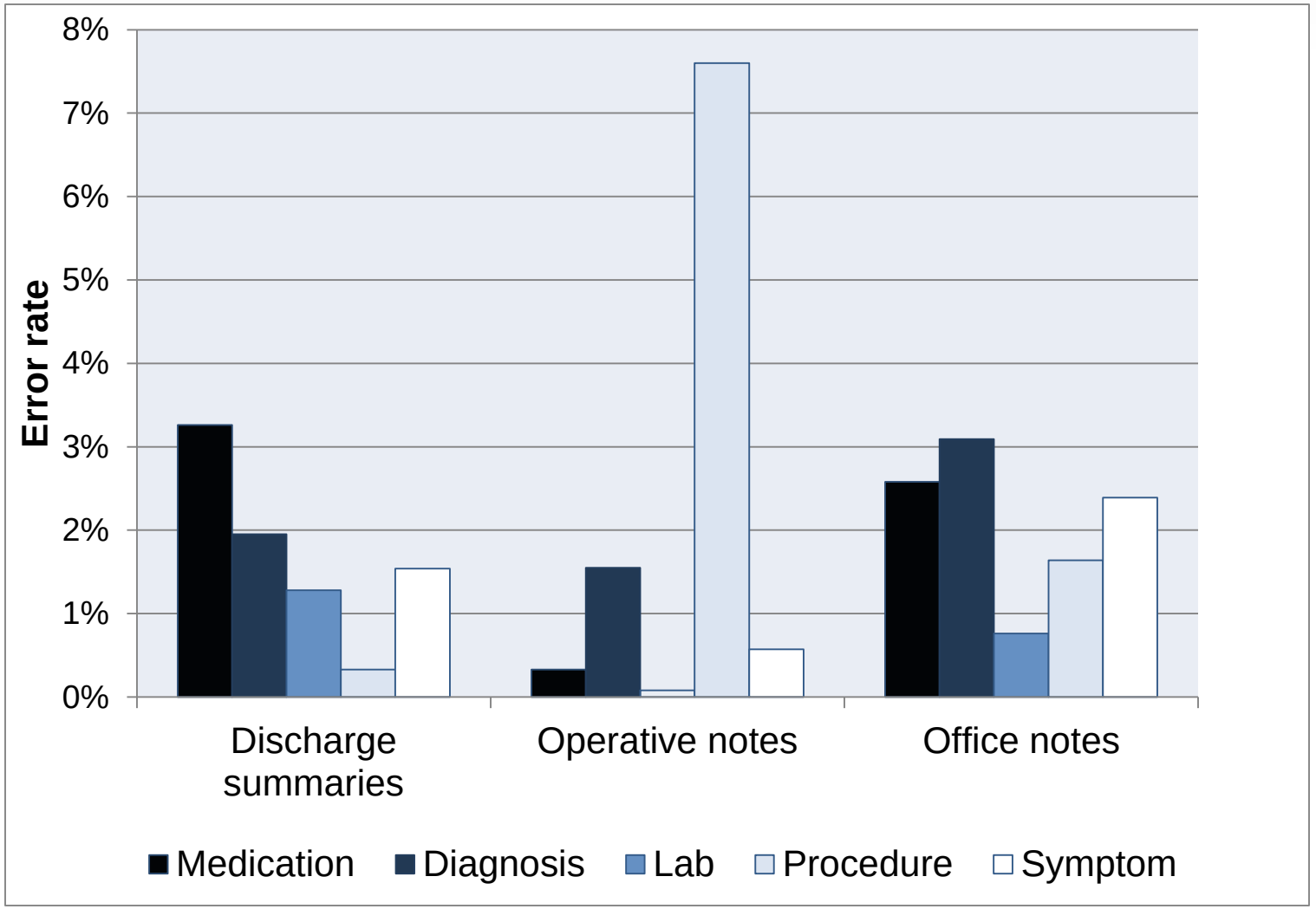


Errors by Semantic Type

		Errors – Semantic Types n (%)[1]							
		General English	Clinical Information						
			Medication	Diagnosis	Procedure	Symptom	Lab	Physical Exam	Imaging
Discharge Summaries (75)	SR	3255 (83.6)	**127 (3.3)**	76 (2.0)	13 (0.3)	60 (1.5)	50 (1.3)	30 (0.8)	25 (0.6)
	MT	124 (63.6)	14 (7.2)	9 (4.6)	2 (1.0)	7 (3.6)	2 (1.0)	14 (7.2)	2 (1.0)
	SN	105 (64.4)	6 (3.7)	9 (5.5)	2 (1.2)	7 (4.3)	2 (1.2)	14 (8.6)	2 (1.2)
Office Notes (55)	SR	1305 (82.2)	41 (2.6)	**49 (3.1)**	26 (1.6)	38 (2.4)	12 (0.8)	7 (0.4)	7 (0.4)
	MT	79 (84.1)	1 (1.0)	4 (4.2)	2 (2.1)	2 (2.1)	1 (1.0)	0 (0.0)	0 (0.0)
	SN	28 (87.5)	0 (0.0)	2 (6.3)	0 (0.0)	1 (3.1)	1 (3.1)	0 (0.0)	0 (0.0)
Operative Notes (39)	SR	947 (77.4)	4 (0.3)	19 (1.6)	**93 (7.6)**	7 (0.6)	1 (0.1)	4 (0.3)	1 (0.1)
	MT	81 (67.5)	0 (0.0)	4 (3.3)	15 (12.5)	2 (1.7)	0 (0.0)	0 (0.0)	1 (0.8)
	SN	72 (75.0)	0 (0.0)	3 (3.1)	9 (9.4)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.0)
All Notes (169)	SR	5507 (82.2)	**172 (2.6)**	144 (2.2)	132 (2.0)	105 (1.6)	63 (0.9)	41 (0.6)	33 (0.5)
	MT	284 (69.1)	**15 (3.7)**	17 (4.1)	19 (4.6)	12 (2.9)	3 (0.7)	15 (3.7)	3 (0.7)
	SN	205 (70.4)	**6 (2.1)**	14 (4.8)	11 (3.8)	9 (3.1)	3 (1.0)	15 (5.2)	3 (1.0)

[1] n = number of errors; % = number of errors of a specific type divided by the total number of errors

Clinical Information Errors in SR Notes



Clinical Information Errors Across Note Stages

	Total Errors	Clinical Information Errors n (%)	General English Errors n (%)	Other Errors n (%)
SR	6703	691 (10.3)	5507 (82.2)	505 (7.5)
MT	411	84 (20.4)	284 (69.1)	43 (10.5)
SN	291	61 (21.0)	205 (70.44)	25 (8.6)

Other errors include patient and provider information, dates and ???.

- 40% of SR notes, 7% of MT notes, and 5% of SN notes contain at least one clinically significant error.

Content Rearranging and Stylistic Changes

	Medical Transcriptionist	Clinician
Rearranged text	17.4%	6.4%
Made stylistic changes	91.7%	43.1%
Added information	N/A	29.7%
Deleted information	N/A	24.8%

% = number of notes where changes were made divided by total number of notes.

Error Examples

- *TB vaccine* vs. *TD vaccine.*
- *Staining* vs. *standing.*
  - "Continues to have daily standing" (pt's menstruation).
  - SR, transcriptionist, and signed note all missed this error.
- *Menorrhagia* and *gluten allergy* were missed by SR and transcriptionist, and remained omitted on the signed note.
- SR and transcriptionist missed the name of the drug and listed as *??__??*
  - The drug was *celecoxib*. The SR and transcriptionist notes did not record it. The signed note listed drug as *naproxen*!

Discussion: Productivity and Quality

- Back-end dictation service had a relatively quick turnaround time and a low error rate.
- While many errors were generated by SR, most (~94%) were corrected by the medical transcriptionist manually.
  - The addition of a human editing an SR-generated note is invaluable.
- 20% of EHR-related malpractice cases were due to incorrect information in the EHR.[1]
  - Without the MT revising the notes, clinically significant errors could have had a negative impact on patient care and potentially caused legal issues.

[1]Ruder DB. Malpractice claims analysis confirms risks in EHRs. Patient Safety & Quality Healthcare. Jan/Feb 2014: Volume 11, Issue 1.

Discussion: Error Checking

- Errors still left in the signed note suggest that some providers may not review their dictated notes thoroughly or at all.
  - 7% of signed notes contained a blank space the transcriptionist marked as *??___??*
- If physicians use SR directly, they may have to spend a considerable amount of time correcting the SR-generated text.
  - Although the errors might be less than our results, since the SR can be trained by the individual physician.
- Automated error detection may help improve the accuracy of dictated documents.

Ongoing and Next Steps

- Conduct error analysis for front-end user of SR.
  - Clinical observations
  - Simulations
- Build a knowledge base.
  - Confusion sets
  - Error frequencies
  - Error patterns
- Develop automated methods to detect SR errors.
  - Statistical methods (noisy channel models, co-occurrence statistics), machine learning, and knowledge-based methods.

Acknowledgements

- This study is funded by AHRQ R01HS024264.
- SR Research Team
  - David Bates
  - David Mack
  - Evgeni Kontrient
  - Foster Goss
  - Jessica Huynh
  - Kenneth Lai
  - Leigh Kowalski
  - Marie Meteer
  - Ray Doan
  - Shiri Hassid
  - Suzanne Blackley
  - Warren Acker
- BWH Dragon Steering Committee, BWH Health Informatics Services, Partners eCare Clinical Informatics Team, and Partners Clinical Quality & Analysis Team

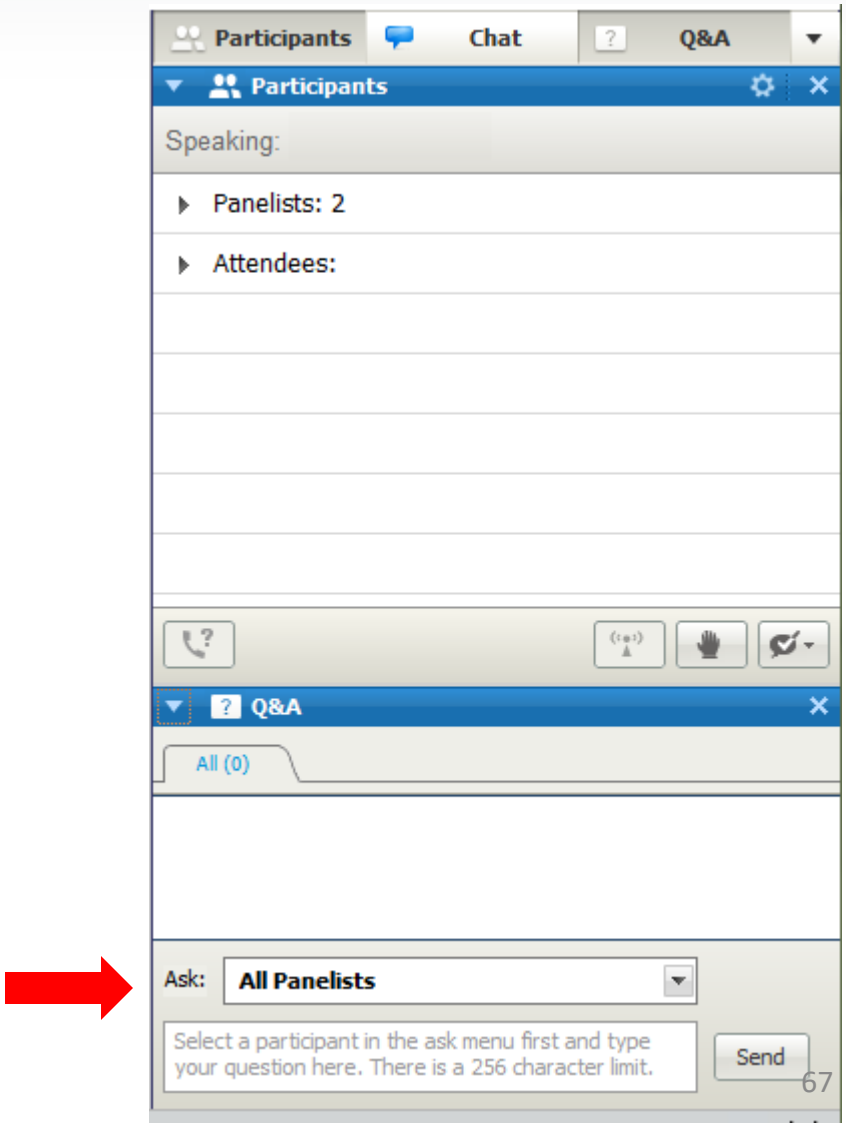
Contact Information

Li Zhou

LZHOU2@PARTNERS.ORG

How to Submit a Question

- At any time during the presentation, type your question into the "Q&A" section of your WebEx Q&A panel.
- Please address your questions to "All Panelists" in the drop-down menu.
- Select "Send" to submit your question to the moderator.
- Questions will be read aloud by the moderator.



Obtaining CME/CE Credits

If you would like to receive continuing education credit for this activity, please visit

http://hitwebinar.cds.pesgce.com/eindex.php